

SEQUENCE LISTING

<110> Genetica, Inc.
Beach, David H.
Hannon, Gregory J.

<120> METHODS AND REAGENTS FOR INCREASING PROLIFERATIVE
CAPACITY AND PREVENTING REPLICATIVE SENESCENCE

<130> GNCA-P01-005

<140> 09/936,035

<141> 1999-11-24

<150> PCT/US99/27907

<151> 1999-11-24

<150> 60/120,549

<151> 1999-02-17

<150> 60/109,891

<151> 1998-11-25

<160> 43

<170> PatentIn Ver. 2.1

<210> 1

<211> 4027

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (57)..(3452)

<400> 1

caggcagcgt ggtcctgctg cgcacgtggg aagccctggc cccggccacc cccgcg atg 59
Met
1

ccg cgc gct ccc cgc tgc cga gcc gtg cgc tcc ctg ctg cgc agc cac 107
Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser His
5 10 15

tac cgc gag gtg ctg ccg ctg gcc acg ttc gtg cgg cgc ctg ggg ccc 155
Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly Pro
20 25 30

cag ggc tgg cgg ctg gtg cag cgc ggg gac ccg gcg gct ttc cgc gcg 203
Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg Ala
35 40 45

ctg gtg gcc cag tgc ctg gtg tgc gtg ccc tgg gac gca cgg ccg ccc 251
Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro Pro
50 55 60 65

ccc gcc gcc ccc tcc ttc cgc cag gtg tcc tgc ctg aag gag ctg gtg	299
Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu Val	
70 75 80	
gcc cga gtg ctg cag agg ctg tgc gag cgc ggc gcg aag aac gtg ctg	347
Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val Leu	
85 90 95	
gcc ttc ggc ttc gcg ctg ctg gac ggg gcc cgc ggg ggc ccc ccc gag	395
Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro Glu	
100 105 110	
gcc ttc acc acc agc gtg cgc agc tac ctg ccc aac acg gtg acc gac	443
Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr Asp	
115 120 125	
gca ctg cgg ggg agc ggg gcg tgg ggg ctg ctg ttg cgc cgc gtg ggc	491
Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val Gly	
130 135 140 145	
gac gac gtg ctg gtt cac ctg ctg gca cgc tgc gcg ctc ttt gtg ctg	539
Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val Leu	
150 155 160	
gtg gct ccc agc tgc gcc tac cag gtg tgc ggg ccg ccg ctg tac cag	587
Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr Gln	
165 170 175	
ctc ggc gct gcc act cag gcc cgg ccc ccg cca cac gct agt gga ccc	635
Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly Pro	
180 185 190	
cga agg cgt ctg gga tgc gaa cgg gcc tgg aac cat agc gtc agg gag	683
Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg Glu	
195 200 205	
gcc ggg gtc ccc ctg ggc ctg cca gcc ccg ggt gcg agg agg cgc ggg	731
Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg Gly	
210 215 220 225	
ggc agt gcc agc cga agt ctg ccg ttg ccc aag agg ccc agg cgt ggc	779
Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg Gly	
230 235 240	
gct gcc cct gag ccg gag cgg acg ccc gtt ggg cag ggg tcc tgg gcc	827
Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp Ala	
245 250 255	
cac ccg ggc agg acg cgt gga ccg agt gac cgt ggt ttc tgt gtg gtg	875
His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val Val	
260 265 270	
tca cct gcc aga ccc gcc gaa gaa gcc acc tct ttg gag ggt gcg ctc	923
Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala Leu	
275 280 285	
tct ggc acg cgc cac tcc cac cca tcc gtg ggc cgc cag cac cac gcg	971

Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His Ala	
290 295 300 305	
ggc ccc cca tcc aca tcg cgg cca cca cgt ccc tgg gac acg cct tgt	1019
Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro Cys	
310 315 320	
ccc ccg gtg tac gcc gag acc aag cac ttc ctc tac tcc tca ggc gac	1067
Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly Asp	
325 330 335	
aag gag cag ctg cgg ccc tcc ttc cta ctc agc tct ctg agg ccc agc	1115
Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro Ser	
340 345 350	
ctg act ggc gct cgg agg ctc gtg gag acc atc ttt ctg ggt tcc agg	1163
Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser Arg	
355 360 365	
ccc tgg atg cca ggg act ccc cgc agg ttg ccc cgc ctg ccc cag cgc	1211
Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln Arg	
370 375 380 385	
tac tgg caa atg cgg ccc ctg ttt ctg gag ctg ctt ggg aac cac gcg	1259
Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His Ala	
390 395 400	
cag tgc ccc tac ggg gtg ctc ctc aag acg cac tgc ccg ctg cga gct	1307
Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg Ala	
405 410 415	
gcg gtc acc cca gca gcc ggt gtc tgt gcc cgg gag aag ccc cag ggc	1355
Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln Gly	
420 425 430	
tct gtg gcg gcc ccc gag gag gag gac aca gac ccc cgt cgc ctg gtg	1403
Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu Val	
435 440 445	
cag ctg ctc cgc cag cac agc agc ccc tgg cag gtg tac ggc ttc gtg	1451
Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe Val	
450 455 460 465	
cgg gcc tgc ctg cgc cgg ctg gtg ccc cca ggc ctc tgg ggc tcc agg	1499
Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser Arg	
470 475 480	
cac aac gaa cgc cgc ttc ctc agg aac acc aag aag ttc atc tcc ctg	1547
His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser Leu	
485 490 495	
ggg aag cat gcc aag ctc tcg ctg cag gag ctg acg tgg aag atg agc	1595
Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met Ser	
500 505 510	
gtg cgg ggc tgc gct tgg ctg cgc agg agc cca ggg gtt ggc tgt gtt	1643
Val Arg Gly Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys Val	

515	520	525	
ccg gcc gca gag cac cgt ctg cgt gag gag atc ctg gcc aag ttc ctg			1691
Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe Leu			
530	535	540	545
cac tgg ctg atg agt gtg tac gtc gtc gag ctg ctc agg tct ttc ttt			1739
His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe			
	550	555	560
tat gtc acg gag acc acg ttt caa aag aac agg ctc ttt ttc tac cgg			1787
Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr Arg			
	565	570	575
aag agt gtc tgg agc aag ttg caa agc att gga atc aga cag cac ttg			1835
Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His Leu			
	580	585	590
aag agg gtg cag ctg cgg gag ctg tcg gaa gca gag gtc agg cag cat			1883
Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln His			
	595	600	605
cgg gaa gcc agg ccc gcc ctg ctg acg tcc aga ctc cgc ttc atc ccc			1931
Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro			
	610	615	620
aag cct gac ggg ctg cgg ccg att gtg aac atg gac tac gtc gtg gga			1979
Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val Gly			
	630	635	640
gcc aga acg ttc cgc aga gaa aag agg gcc gag cgt ctc acc tcg agg			2027
Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser Arg			
	645	650	655
gtg aag gca ctg ttc agc gtg ctc aac tac gag cgg gcg cgg cgc ccc			2075
Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg Pro			
	660	665	670
ggc ctc ctg ggc gcc tct gtg ctg ggc ctg gac gat atc cac agg gcc			2123
Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala			
	675	680	685
tgg cgc acc ttc gtg ctg cgt gtg cgg gcc cag gac ccg ccg cct gag			2171
Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro Glu			
	690	695	700
ctg tac ttt gtc aag gtg gat gtg acg ggc gcg tac gac acc atc ccc			2219
Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile Pro			
	710	715	720
cag gac agg ctc acg gag gtc atc gcc agc atc atc aaa ccc cag aac			2267
Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln Asn			
	725	730	735
acg tac tgc gtg cgt cgg tat gcc gtg gtc cag aag gcc gcc cat ggg			2315
Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His Gly			
	740	745	750

cac gtc cgc aag gcc ttc aag agc cac gtc tct acc ttg aca gac ctc His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp Leu 755 760 765	2363
cag ccg tac atg cga cag ttc gtg gct cac ctg cag gag acc agc ccg Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser Pro 770 775 780 785	2411
ctg agg gat gcc gtc gtc atc gag cag agc tcc tcc ctg aat gag gcc Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu Ala 790 795 800	2459
agc agt ggc ctc ttc gac gtc ttc cta cgc ttc atg tgc cac cac gcc Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His Ala 805 810 815	2507
gtg cgc atc agg ggc aag tcc tac gtc cag tgc cag ggg atc ccg cag Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro Gln 820 825 830	2555
ggc tcc atc ctc tcc acg ctg ctc tgc agc ctg tgc tac ggc gac atg Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp Met 835 840 845	2603
gag aac aag ctg ttt gcg ggg att cgg cgg gac ggg ctg ctc ctg cgt Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu Arg 850 855 860 865	2651
ttg gtg gat gat ttc ttg ttg gtg aca cct cac ctc acc cac gcg aaa Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala Lys 870 875 880	2699
acc ttc ctc agg acc ctg gtc cga ggt gtc cct gag tat ggc tgc gtg Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys Val 885 890 895	2747
gtg aac ttg cgg aag aca gtg gtg aac ttc cct gta gaa gac gag gcc Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu Ala 900 905 910	2795
ctg ggt ggc acg gct ttt gtt cag atg ccg gcc cac ggc cta ttc ccc Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe Pro 915 920 925	2843
tgg tgc ggc ctg ctg ctg gat acc cgg acc ctg gag gtg cag agc gac Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser Asp 930 935 940 945	2891
tac tcc agc tat gcc cgg acc tcc atc aga gcc agt ctc acc ttc aac Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe Asn 950 955 960	2939
cgc ggc ttc aag gct ggg agg aac atg cgt cgc aaa ctc ttt ggg gtc Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly Val 965 970 975	2987

ttg cgg ctg aag tgt cac agc ctg ttt ctg gat ttg cag gtg aac agc 3035
 Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn Ser
 980 985 990

ctc cag acg gtg tgc acc aac atc tac aag atc ctc ctg ctg cag gcg 3083
 Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln Ala
 995 1000 1005

tac agg ttt cac gca tgt gtg ctg cag ctc cca ttt cat cag caa gtt 3131
 Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln Val
 1010 1015 1020 1025

tgg aag aac ccc aca ttt ttc ctg cgc gtc atc tct gac acg gcc tcc 3179
 Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala Ser
 1030 1035 1040

ctc tgc tac tcc atc ctg aaa gcc aag aac gca ggg atg tgc ctg ggg 3227
 Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu Gly
 1045 1050 1055

gcc aag ggc gcc gcc ggc cct ctg ccc tcc gag gcc gtg cag tgg ctg 3275
 Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp Leu
 1060 1065 1070

tgc cac caa gca ttc ctg ctc aag ctg act cga cac cgt gtc acc tac 3323
 Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr Tyr
 1075 1080 1085

gtg cca ctc ctg ggg tca ctc agg aca gcc cag acg cag ctg agt cgg 3371
 Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser Arg
 1090 1095 1100 1105

aag ctc ccg ggg acg acg ctg act gcc ctg gag gcc gca gcc aac ccg 3419
 Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn Pro
 1110 1115 1120

gca ctg ccc tca gac ttc aag acc atc ctg gac tgatggccac ccgcccacag 3472
 Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp
 1125 1130

ccaggccgag agcagacacc agcagccctg tcacgccggg ctctacgtcc cagggaggga 3532

ggggcgggcc acaccaggc ccgcaccgct gggagtctga ggcctgagtg agtgtttggc 3592

cgaggcctgc atgtccggct gaaggctgag tgtccggctg aggcctgagc gagtgtccag 3652

ccaagggctg agtgtccagc acacctgccg tcttcacttc cccacagget ggcgctcggc 3712

tccaccccag ggccagcttt tctcaccag gagcccggct tccactcccc acataggaat 3772

agtcctatccc cagattcgcc attgttcacc cctcgccctg ccttcctttg ccttccaccc 3832

ccaccatcca ggtggagacc ctgagaagga ccctgggagc tctgggaatt tggagtgacc 3892

aaaggtgtgc cctgtacaca ggcgaggacc ctgcacctgg atgggggtcc ctgtgggtca 3952

aattgggggg aggtgctgtg ggagtaaaat actgaatata tgagtttttc agttttgaaa 4012

aaaaaaaaaa aaaaa

4027

<210> 2
<211> 1132
<212> PRT
<213> Homo sapiens

<400> 2
Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser
1 5 10 15
His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly
20 25 30
Pro Gln Gly Trp Arg Leu Val Gln Arg Gly Asp Pro Ala Ala Phe Arg
35 40 45
Ala Leu Val Ala Gln Cys Leu Val Cys Val Pro Trp Asp Ala Arg Pro
50 55 60
Pro Pro Ala Ala Pro Ser Phe Arg Gln Val Ser Cys Leu Lys Glu Leu
65 70 75 80
Val Ala Arg Val Leu Gln Arg Leu Cys Glu Arg Gly Ala Lys Asn Val
85 90 95
Leu Ala Phe Gly Phe Ala Leu Leu Asp Gly Ala Arg Gly Gly Pro Pro
100 105 110
Glu Ala Phe Thr Thr Ser Val Arg Ser Tyr Leu Pro Asn Thr Val Thr
115 120 125
Asp Ala Leu Arg Gly Ser Gly Ala Trp Gly Leu Leu Leu Arg Arg Val
130 135 140
Gly Asp Asp Val Leu Val His Leu Leu Ala Arg Cys Ala Leu Phe Val
145 150 155 160
Leu Val Ala Pro Ser Cys Ala Tyr Gln Val Cys Gly Pro Pro Leu Tyr
165 170 175
Gln Leu Gly Ala Ala Thr Gln Ala Arg Pro Pro Pro His Ala Ser Gly
180 185 190
Pro Arg Arg Arg Leu Gly Cys Glu Arg Ala Trp Asn His Ser Val Arg
195 200 205
Glu Ala Gly Val Pro Leu Gly Leu Pro Ala Pro Gly Ala Arg Arg Arg
210 215 220
Gly Gly Ser Ala Ser Arg Ser Leu Pro Leu Pro Lys Arg Pro Arg Arg
225 230 235 240
Gly Ala Ala Pro Glu Pro Glu Arg Thr Pro Val Gly Gln Gly Ser Trp
245 250 255

Ala His Pro Gly Arg Thr Arg Gly Pro Ser Asp Arg Gly Phe Cys Val
 260 265 270
 Val Ser Pro Ala Arg Pro Ala Glu Glu Ala Thr Ser Leu Glu Gly Ala
 275 280 285
 Leu Ser Gly Thr Arg His Ser His Pro Ser Val Gly Arg Gln His His
 290 295 300
 Ala Gly Pro Pro Ser Thr Ser Arg Pro Pro Arg Pro Trp Asp Thr Pro
 305 310 315 320
 Cys Pro Pro Val Tyr Ala Glu Thr Lys His Phe Leu Tyr Ser Ser Gly
 325 330 335
 Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro
 340 345 350
 Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser
 355 360 365
 Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln
 370 375 380
 Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His
 385 390 395 400
 Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg
 405 410 415
 Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln
 420 425 430
 Gly Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu
 435 440 445
 Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe
 450 455 460
 Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser
 465 470 475 480
 Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser
 485 490 495
 Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met
 500 505 510
 Ser Val Arg Gly Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys
 515 520 525
 Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe
 530 535 540
 Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe
 545 550 555 560

Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr
565 570 575
Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His
580 585 590
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln
595 600 605
His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile
610 615 620
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val
625 630 635 640
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser
645 650 655
Arg Val Lys Ala Leu Phe Ser Val Leu Asn Tyr Glu Arg Ala Arg Arg
660 665 670
Pro Gly Leu Leu Gly Ala Ser Val Leu Gly Leu Asp Asp Ile His Arg
675 680 685
Ala Trp Arg Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro
690 695 700
Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr Asp Thr Ile
705 710 715 720
Pro Gln Asp Arg Leu Thr Glu Val Ile Ala Ser Ile Ile Lys Pro Gln
725 730 735
Asn Thr Tyr Cys Val Arg Arg Tyr Ala Val Val Gln Lys Ala Ala His
740 745 750
Gly His Val Arg Lys Ala Phe Lys Ser His Val Ser Thr Leu Thr Asp
755 760 765
Leu Gln Pro Tyr Met Arg Gln Phe Val Ala His Leu Gln Glu Thr Ser
770 775 780
Pro Leu Arg Asp Ala Val Val Ile Glu Gln Ser Ser Ser Leu Asn Glu
785 790 795 800
Ala Ser Ser Gly Leu Phe Asp Val Phe Leu Arg Phe Met Cys His His
805 810 815
Ala Val Arg Ile Arg Gly Lys Ser Tyr Val Gln Cys Gln Gly Ile Pro
820 825 830
Gln Gly Ser Ile Leu Ser Thr Leu Leu Cys Ser Leu Cys Tyr Gly Asp
835 840 845
Met Glu Asn Lys Leu Phe Ala Gly Ile Arg Arg Asp Gly Leu Leu Leu
850 855 860

Arg Leu Val Asp Asp Phe Leu Leu Val Thr Pro His Leu Thr His Ala
 865 870 875 880
 Lys Thr Phe Leu Arg Thr Leu Val Arg Gly Val Pro Glu Tyr Gly Cys
 885 890 895
 Val Val Asn Leu Arg Lys Thr Val Val Asn Phe Pro Val Glu Asp Glu
 900 905 910
 Ala Leu Gly Gly Thr Ala Phe Val Gln Met Pro Ala His Gly Leu Phe
 915 920 925
 Pro Trp Cys Gly Leu Leu Leu Asp Thr Arg Thr Leu Glu Val Gln Ser
 930 935 940
 Asp Tyr Ser Ser Tyr Ala Arg Thr Ser Ile Arg Ala Ser Leu Thr Phe
 945 950 955 960
 Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly
 965 970 975
 Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn
 980 985 990
 Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln
 995 1000 1005
 Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln Gln
 1010 1015 1020
 Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp Thr Ala
 1025 1030 1035 1040
 Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly Met Ser Leu
 1045 1050 1055
 Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu Ala Val Gln Trp
 1060 1065 1070
 Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr Arg His Arg Val Thr
 1075 1080 1085
 Tyr Val Pro Leu Leu Gly Ser Leu Arg Thr Ala Gln Thr Gln Leu Ser
 1090 1095 1100
 Arg Lys Leu Pro Gly Thr Thr Leu Thr Ala Leu Glu Ala Ala Ala Asn
 1105 1110 1115 1120
 Pro Ala Leu Pro Ser Asp Phe Lys Thr Ile Leu Asp
 1125 1130

<210> 3
 <211> 2372
 <212> DNA

<213> Homo sapiens

<400> 3

```
gcaccgcgcg agcttggctg cttctggggc ctgtgtggcc ctgtgtgtcg gaaagatgga 60
gcaagaagcc gagcccgagg ggccggccgcg acccctctga ccgagatcct gctgctttcg 120
cagccaggag caccgtccct ccccgatta gtgcgtacga gcgcccagtg ccctggcccg 180
gagagtggaa tgatccccga ggcccagggc gtcgtgcttc cgcagtagtc agtccccgtg 240
aaggaaactg gggagtcttg agggaccccc gactccaagc gcgaaaaccc cggaatggtga 300
ggagcaggca aatgtgcaat accaacaatgt ctgtacctac tgatggtgct gtaaccacct 360
cacagattcc agcttcggaa caagagaccc tggtagacc aaagccattg cttttgaagt 420
tattaaagtc tgttggtgca caaaaagaca cttatactat gaaagagggt cttttttatc 480
ttggccagta tattatgact aaacgattat atgatgagaa gcaacaacat attgtatatt 540
gttcaaata tcttctagga gatttgtttg gcgtgccaaag cttctctgtg aaagagcaca 600
ggaaaatata taccatgatc tacaggaaact tggtagtagt caatcagcag gaatcatcgg 660
actcaggtac atctgtgagt gagaacaggt gtcaccttga aggtgggagt gatcaaaagg 720
acctgtgaca agagcttcag gaagagaaac cttcatcttc acatttggtt tctagaccat 780
ctacctcatc tagaaggaga gcaattagt agacagaaga aaattcagat gaattatctg 840
gtgaacgaca aagaaaacgc cacaatatct atagtatttc ctttctcttt gatgaaagcc 900
tggctctgtg tgtaataagg gagatatgtt gtgaaagaag cagtagcagt gaatctacag 960
ggacgccatc gaatccggat cttgatgctg gtgtaagtga acattcaggt gattgggttg 1020
atcaggattc agtttcagat cagtttagtg tagaatttga agttgaatct ctcgactcag 1080
aagattatag ccttagtgaa gaaggacaag aactctcaga tgaagatgat gaggtatatc 1140
aagttactgt gtatcaggca ggggagagt atacagattc atttgaagaa gatcctgaaa 1200
tttccttagc tgactatttg aaatgcactt catgcaatga aatgaatccc ccccttccat 1260
cacattgcaa cagatgttgg gcccttcgtg agaattggct tcctgaagat aaagggaaaag 1320
ataaagggga aatctctgag aaagccaaac tggaaaactc aacacaagct gaagagggct 1380
ttgatgttcc tgattgtaaa aaaactatag tgaatgattc cagagagtca tgtgttgagg 1440
aaaatgatga taaaattaca caagcttcac aatcacaaga aagtgaagac tattctcagc 1500
catcaacttc tagtagcatt atttatagca gccaaagaaga tgtgaaagag tttgaaaggg 1560
aagaaaccca agacaaagaa gagagtgtgg aatctagttt gccccttaat gccattgaac 1620
cttgtgtgat ttgtcaaggt cgacctaaaa atggttgcat tgtccatggc aaaacaggac 1680
atcttatggc ctgctttaca tgtgcaaaga agctaaagaa aaggaataag ccctgcccag 1740
tatgtagaca accaattcaa atgattgtgc taacttattt cccctagttg acctgtctat 1800
aagagaatta tatatttcta actatataac ctaggaatt tagacaacct gaaatttatt 1860
cacatatatc aaagtgaaga aatgcctcaa ttcacataga tttcttctct ttagtataat 1920
tgacctactt tggtagtgga atagtgaata ctactataa tttgacttga atatgtagct 1980
catcctttac accaactcct aattttaaat aatttctact ctgtcttaaa tgagaagtac 2040
ttggtttttt ttttcttaaa tatgtatatg acatttaaat gtaacttatt attttttttg 2100
agaccgagtc ttgctctgtt acccaggctg gagtgcagtg ggtgatcttg gctcactgca 2160
agctctgccc tccccgggtt cgcaccattc tcctgcctca gcctcccaat tagcttggcc 2220
tacagtcatc tgccaccaca cctggctaata tttttgtact tttagtagag acagggtttc 2280
accgtgttag ccaggatggt ctcgatctcc tgacctcgtg atccgcccac ctcggcctcc 2340
caaagtgcgt ggattacagg catgagccac cg 2372
```

<210> 4

<211> 491

<212> PRT

<213> Homo sapiens

<400> 4

```
Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr
  1                      5                      10                      15
Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
                20                      25                      30
```

Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys Asp Thr Tyr
 35 40 45
 Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile Met Thr Lys
 50 55 60
 Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys Ser Asn Asp
 65 70 75 80
 Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
 85 90 95
 Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Gln
 100 105 110
 Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His
 115 120 125
 Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu Leu Gln Glu
 130 135 140
 Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Ser Ser
 145 150 155 160
 Arg Arg Arg Ala Ile Ser Glu Thr Glu Glu Asn Ser Asp Glu Leu Ser
 165 170 175
 Gly Glu Arg Gln Arg Lys Arg His Lys Ser Asp Ser Ile Ser Leu Ser
 180 185 190
 Phe Asp Glu Ser Leu Ala Leu Cys Val Ile Arg Glu Ile Cys Cys Glu
 195 200 205
 Arg Ser Ser Ser Ser Glu Ser Thr Gly Thr Pro Ser Asn Pro Asp Leu
 210 215 220
 Asp Ala Gly Val Ser Glu His Ser Gly Asp Trp Leu Asp Gln Asp Ser
 225 230 235 240
 Val Ser Asp Gln Phe Ser Val Glu Phe Glu Val Glu Ser Leu Asp Ser
 245 250 255
 Glu Asp Tyr Ser Leu Ser Glu Glu Gly Gln Glu Leu Ser Asp Glu Asp
 260 265 270
 Asp Glu Val Tyr Gln Val Thr Val Tyr Gln Ala Gly Glu Ser Asp Thr
 275 280 285
 Asp Ser Phe Glu Glu Asp Pro Glu Ile Ser Leu Ala Asp Tyr Trp Lys
 290 295 300
 Cys Thr Ser Cys Asn Glu Met Asn Pro Pro Leu Pro Ser His Cys Asn
 305 310 315 320
 Arg Cys Trp Ala Leu Arg Glu Asn Trp Leu Pro Glu Asp Lys Gly Lys
 325 330 335

Asp Lys Gly Glu Ile Ser Glu Lys Ala Lys Leu Glu Asn Ser Thr Gln
340 345 350

Ala Glu Glu Gly Phe Asp Val Pro Asp Cys Lys Lys Thr Ile Val Asn
355 360 365

Asp Ser Arg Glu Ser Cys Val Glu Glu Asn Asp Asp Lys Ile Thr Gln
370 375 380

Ala Ser Gln Ser Gln Glu Ser Glu Asp Tyr Ser Gln Pro Ser Thr Ser
385 390 395 400

Ser Ser Ile Ile Tyr Ser Ser Gln Glu Asp Val Lys Glu Phe Glu Arg
405 410 415

Glu Glu Thr Gln Asp Lys Glu Glu Ser Val Glu Ser Ser Leu Pro Leu
420 425 430

Asn Ala Ile Glu Pro Cys Val Ile Cys Gln Gly Arg Pro Lys Asn Gly
435 440 445

Cys Ile Val His Gly Lys Thr Gly His Leu Met Ala Cys Phe Thr Cys
450 455 460

Ala Lys Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro Val Cys Arg Gln
465 470 475 480

Pro Ile Gln Met Ile Val Leu Thr Tyr Phe Pro
485 490

<210> 5

<211> 4233

<212> DNA

<213> Homo sapiens

<220>

<223> "n" at various positions may be a, t, c, g, other or unknown

<400> 5

ccctcctccc agtcgaagca cctcctgtcc gccctcagc gcatgggtgg cggtcacgtg 60
cccagaacgt ccggcggtcg ccccgccctc ccagtttccg cgcgcctctt tggcagctgg 120
tcacatggtg aggggtgggg tgagggggcc tctctagctt gcggcctgtg tctatggctg 180
ggccctctgc gtccagctgc tccggaccga gctcgggtgt atggggccgt aggaaccggc 240
tccggggccc cgataacggg ccgccccccac agcaccgccg gctggcggtga ggtaagtga 300
gtcccttccc aggaatgaga accagtgcgc gcccccctca cagctttcca cgcgttcgtt 360
tcgcgagctg gttatggaag ggtcgctcaa gggcggaag tggggccttt gtggtcatgg 420
gaaagtataa ttttagggac tgaggtgtag gatcttcgat gcaaggcatg tgtcatgtgt 480
gatctttgtg cggggcgcgga ttgtcccaaa ggaaaaagcg ttttctattg cagggcctca 540
cgtggctgga ggggttggtg ttgagtcatt gtgttatctc tggggccggc cccaaggaag 600
actgggagcg ggggatggga tgctgggtgt gttctttgcg cttttttttt gggagtcctt 660
ttgttgctgc aggtcatacc atcctaactc tgtaagcgac ttttggtgat aggagtctgt 720
gattgtaggg tctcccttga totgagaatg gctacctctc gatatgagcc agtggtgaa 780
attggtgtcg gtgcctatgg gacagtgtac aaggcccgtg atccccacag tggccacttt 840
gtggccctca agagtgtgag agtccccaat ggaggaggag gtggaggagg ccttcccatc 900
agcacagttc gtgaggtggc ttactgagg cgactggagg cttttgagca tcccaatgtt 960
gtccggtgag aagggtggtg aggggtgggc gtggggagta aagggaaaag acagcctata 1020

ggtggggtgt	gatgatctgt	agagaagtgg	ggaccctgag	gaaataatga	gaggccatgt	1080
tgggttaaag	gggattgaaa	agtgagcatt	tactctggtc	aggctgatgg	acgtctgtgc	1140
cacatcccga	actgaccggg	agatcaaggt	aaccctgggtg	tttgagcatg	tagaccagga	1200
cctaaggaca	tatctggaca	aggcaccccc	accaggcttg	ccagccgaaa	cgatcaaggt	1260
gagtgggggtt	ggtaggcatt	gagaggtgga	ttgggacctt	tgtagtagaa	ccttctggga	1320
tttcagggtat	ggtgcctagt	ttccagtgca	tctgtacctc	cccctttgaa	actaggatct	1380
gatgcgccag	tttctaagag	gcctagattt	ccttcatgcc	aattgcatcg	ttcaccgaga	1440
tctgaagcca	gagaacattc	tggtgacaag	tggtggaaca	gtcaagctgg	ctgactttgg	1500
cctggccaga	atctacagct	accagatggc	acttacaccc	gtggtcagta	gaaagatggg	1560
acaaaaatgg	gttctggttg	ggaataggag	agtgattgcc	cgtagcaatt	gagaagtcac	1620
gtgcttcatg	tgttcagtc	agcaagttgt	gtttcatggg	aaccatggg	gtccccatcc	1680
attcttccta	ttccctttag	gttggttacac	tctggtaccg	agctcccgaa	gttcttctgc	1740
agtcacata	tgcaacacct	gtggacatgt	ggagtgttgg	ctgtatcttt	gcagagatgt	1800
ttcgtcgaaa	gtatgggacc	cacataccct	ggactacctt	gaattcccca	aatcgcttgt	1860
tcataaacca	catccatacc	ttgccattc	tttttttttg	agaccagggc	ttgctgtgtt	1920
gcccaggctg	gattgcaatg	gcatgatcac	agctcactgc	agcttcaacc	tcctgggctc	1980
aagtgatcct	cccatctcag	cttcccaact	agctgacact	acaggcacgc	acctccatgc	2040
ttggctagtt	tgtaaatatt	tttatagaga	tggggtctca	gtatattgcc	caggctgggc	2100
ttgaactctt	gcactcaagc	aatcctccca	cccctacctc	ccaaagtagc	ataagctact	2160
gcatctggcc	ccattctttt	acttgcgtag	tactaacttg	cccatagcag	aaagctctga	2220
aatgttctgg	aattaggaac	ttcatatccc	tttattctct	ttatttttta	tttattttatt	2280
tattttattta	tttattttatt	gagataaggt	ttcactctgn	naccagggct	ggagtncagt	2340
ggcccaatta	nagctcactg	tancctctac	ctcctgggct	aaagmaatcc	tcccatctca	2400
gccccttgag	tancctgagac	taaaggtgca	cgccaccatg	actggctttt	ttttttttta	2460
gatggagtct	tgctctgtcg	ccaggctgga	gtgcagtagt	gcgatctctg	ctcactgcaa	2520
cctccacctc	ccagattcaa	gcaattctct	tgactcagcc	tcccaagtag	ctgggaccac	2580
aggtgcacgc	caccatgctc	agctaatttt	tgtactttta	gtaatgacag	gtttcaccat	2640
gttgccagg	atggtctcga	tctcttgacc	tcatgatcca	cccacatcag	actcccaaag	2700
tgctaggatt	acaggcgtga	gcnnnngcac	ctggcatttc	ttttttttta	aaaaaagaga	2760
caaggtcttg	cttgcccagg	ctgatctaga	actcctgggc	tcaagcagtc	ctctcacctc	2820
agcatcccaa	agtgtctgaa	ttgttgccct	ttattcccta	tacttcctat	tttgagccac	2880
taagcagtaa	ccattcaact	aagatatctt	tgaaaatgac	tgctacctta	tatcccttct	2940
caccttaggc	ctctcttctg	tggaaactct	gaagccgacc	agttgggcaa	aatctttgag	3000
taagtgacca	acatggggaga	aaaagatttt	ctattctgag	tcctctttct	gctgaacca	3060
ggatggcaac	tggctctgcc	atggggatgg	gaactggagg	accctcctga	ccagagttct	3120
cctgtcccc	acagcctgat	tgggctgcct	ccagaggatg	actggcctcg	agatgtatcc	3180
ctgccccgtg	gagcctttcc	ccccagaggg	ccccgcccag	tgcagtcggg	ggtacctgag	3240
atggaggagt	cgggagcaca	gctgctgctg	gtaactggag	atggctgtgg	gcacagggaa	3300
agaaatagag	actggggaaa	gaaatagagc	agtatgcagg	gccctggcca	ctgtgggttaa	3360
tgaaacttgg	ttggtagatg	gtctgtagt	tttattacag	ctgcaaatag	ccaccacacag	3420
agaaggatat	agaagagAAC	ccatcctggc	tgggcacggg	ggctcacgcc	tgtaatccca	3480
gcactttggg	aggccaaggt	gggcgtatca	cctgagggtca	ggagttcgag	accagcctgg	3540
ccaacatggg	gaaacctcgt	ctctactaaa	agtacaaaaa	taagccgggg	gtggtggcac	3600
acgcctgtaa	tctcagctac	ttgggaggct	gagataggag	aatcacttca	actcaggagg	3660
cggagggttg	agtgagctga	gatcatacca	ttggcactcc	agcctgggtg	atagagcgag	3720
actccgtctn	caaaaaaaaa	aaaaaagaaa	aaagaagaaa	gctcatccca	ggtattgttg	3780
tgggtggcag	aagctgtttt	cttcatgggt	ttctgacctt	tgctctctcc	ctcaggaaat	3840
gctgactttt	aaccacacac	agcgaatctc	tgctttctga	gctctgcagc	actcttatct	3900
acataaggat	gaaggtaatc	cggagtgcgc	aatggagtgg	ctgccatgga	aggaagaaaa	3960
gctgccattt	cccttctgga	cactgagagg	gcaatctttg	cccttatctc	tgaggctatg	4020
gagggtcctc	ctccatcttt	ctacagagat	tactttgctg	ccttaatgac	attccctcc	4080
cacctctcct	tttgaggctt	ctccttctcc	ttcccatctt	tctacactaa	ggggtatgtt	4140
ccctcttgct	cctttcccta	cctttatatt	tggggtcctt	ttttatacag	gaaaaacaaa	4200
accaaagaa	awaatggccc	tttttttttt	ttt			4233

<211> 303

<212> PRT

<213> Homo sapiens

<400> 6

Met	Ala	Thr	Ser	Arg	Tyr	Glu	Pro	Val	Ala	Glu	Ile	Gly	Val	Gly	Ala	
1				5					10					15		
Tyr	Gly	Thr	Val	Tyr	Lys	Ala	Arg	Asp	Pro	His	Ser	Gly	His	Phe	Val	
			20					25					30			
Ala	Leu	Lys	Ser	Val	Arg	Val	Pro	Asn	Gly	Gly	Gly	Gly	Gly	Gly	Gly	
		35					40					45				
Leu	Pro	Ile	Ser	Thr	Val	Arg	Glu	Val	Ala	Leu	Leu	Arg	Arg	Leu	Glu	
	50					55					60					
Ala	Phe	Glu	His	Pro	Asn	Val	Val	Arg	Leu	Met	Asp	Val	Cys	Ala	Thr	
65					70					75					80	
Ser	Arg	Thr	Asp	Arg	Glu	Ile	Lys	Val	Thr	Leu	Val	Phe	Glu	His	Val	
				85					90					95		
Asp	Gln	Asp	Leu	Arg	Thr	Tyr	Leu	Asp	Lys	Ala	Pro	Pro	Pro	Gly	Leu	
			100					105						110		
Pro	Ala	Glu	Thr	Ile	Lys	Asp	Leu	Met	Arg	Gln	Phe	Leu	Arg	Gly	Leu	
		115					120						125			
Asp	Phe	Leu	His	Ala	Asn	Cys	Ile	Val	His	Arg	Asp	Leu	Lys	Pro	Glu	
	130					135					140					
Asn	Ile	Leu	Val	Thr	Ser	Gly	Gly	Thr	Val	Lys	Leu	Ala	Asp	Phe	Gly	
145					150					155					160	
Leu	Ala	Arg	Ile	Tyr	Ser	Tyr	Gln	Met	Ala	Leu	Thr	Pro	Val	Val	Val	
			165						170					175		
Thr	Leu	Trp	Tyr	Arg	Ala	Pro	Glu	Val	Leu	Leu	Gln	Ser	Thr	Tyr	Ala	
			180					185						190		
Thr	Pro	Val	Asp	Met	Trp	Ser	Val	Gly	Cys	Ile	Phe	Ala	Glu	Met	Phe	
		195					200					205				
Arg	Arg	Lys	Pro	Leu	Phe	Cys	Gly	Asn	Ser	Glu	Ala	Asp	Gln	Leu	Gly	
	210					215					220					
Lys	Ile	Phe	Asp	Leu	Ile	Gly	Leu	Pro	Pro	Glu	Asp	Asp	Trp	Pro	Arg	
225				230						235				240		
Asp	Val	Ser	Leu	Pro	Arg	Gly	Ala	Phe	Pro	Pro	Arg	Gly	Pro	Arg	Pro	
			245						250					255		
Val	Gln	Ser	Val	Val	Pro	Glu	Met	Glu	Glu	Ser	Gly	Ala	Gln	Leu	Leu	
			260					265					270			
Leu	Glu	Met	Leu	Thr	Phe	Asn	Pro	His	Lys	Arg	Ile	Ser	Ala	Phe	Arg	

275

280

285

Ala Leu Gln His Ser Tyr Leu His Lys Asp Glu Gly Asn Pro Glu
 290 295 300

<210> 7

<211> 105

<212> PRT

<213> Homo sapiens

<400> 7

Met His Gly Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu Glu
 1 5 10 15

Pro Gln Asn Glu Ile Pro Val Asp Leu Leu Cys His Glu Gln Leu Ser
 20 25 30

Asp Ser Glu Glu Glu Asn Asp Glu Ile Asp Gly Val Asn His Gln His
 35 40 45

Leu Pro Ala Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys Met
 50 55 60

Cys Cys Lys Cys Glu Ala Arg Ile Glu Leu Val Val Glu Ser Ser Ala
 65 70 75 80

Asp Asp Leu Arg Ala Phe Gln Gln Leu Phe Leu Asn Thr Leu Ser Phe
 85 90 95

Val Cys Pro Trp Cys Ala Ser Gln Gln
 100 105

<210> 8

<211> 4839

<212> DNA

<213> Homo sapiens

<400> 8

tccggttttt ctcaggggac gttgaaatta tttttgtaac gggagtcggg agaggacggg 60
 gcgtgccccg cgtgcgcgcg cgtcgtcttc cccggcgctc ctccacagct cgctggctcc 120
 cgccgcggaa aggcgtcatg ccgcccacaa cccccgaaa aacggccgcc accgccgccg 180
 ctgccgcgcg ggaacccccg gcaccgcgcg cgccgcccc tcttgaggag gaccagagc 240
 aggacagcgg cccggaggac ctgcctctcg tcaggcttga gtttgaagaa acagaagaac 300
 ctgattttac tgcattatgt cagaaattaa agataccaga tcatgtcaga gagagagctt 360
 gggttaacttg ggagaaagt tcatctgttg atggagtatt gggaggttat attcaaaaga 420
 aaaaggaact gtggggaatc tgtatcttta ttgcagcagt tgacctagat gagatgtcgt 480
 tcacttttac tgagctacag aaaaacatag aaatcagtgt ccataaattc tttaacttac 540
 taaaagaaat tgataccagt accaaagtgt ataatgctat gtcaagactg ttgaagaagt 600
 atgatgtatt gtttgcactc ttcagcaa attgaaaggac atgtgaactt atatatttga 660
 cacaacccag cagttcgata tctactgaaa taaattctgc attggtgcta aaagtttctt 720
 ggatcacatt tttattagct aaagggaag tattacaaat ggaagatgat ctggtgattt 780
 catttcagtt aatgctatgt gtccttgact attttattaa actctcacct cccatgttgc 840
 tcaaagaacc atataaaaaca gctgttatac ccattaatgg ttcacctcga acaccaggc 900
 gaggtcagaa caggagtgcg cggatagcaa aacaactaga aaatgatata agaattattg 960
 aagttctctg taaagaacat gaatgtaata tagatgaggt gaaaaatggt tatttcaaaa 1020

at tt t t a t a c c	t t t t a t g a a t	t c t c t t g g a c	t t g t a a c a t c	t a a t g g a c t t	c c a g a g g t t g	1080
a a a a t c t t t c	t a a a c g a t a c	g a a g a a a t t t	a t c t t a a a a a	t a a a g a t c t a	g a t g c a a g a t	1140
t a t t t t t g g a	t c a t g a t a a a	a c t c t t c a g a	c t g a t t c t a t	a g a c a g t t t t	g a a a c a c a g a	1200
g a a c a c c a c g	a a a a a g t a a c	c t t g a t g a a g	a g g t g a a t g t	a a t t c c t c c a	c a c a c t c c a g	1260
t t a g g a c t g t	t a t g a a c a c t	a t c c a a c a a t	t a a t g a t g a t	t t t a a a t t c a	g c a a g t g a t c	1320
a a c c t t c a g a	a a a t c t g a t t	t c c t a t t t t a	a c a a c t g c a c	a g t g a a t c c a	a a a g a a a g t a	1380
t a c t g a a a a g	a g t g a a g g a t	a t a g g a t a c a	t c t t t a a a g a	g a a a t t t g c t	a a a g c t g t g g	1440
g a c a g g g t t g	t g t c g a a a t t	g g a t c a c a g c	g a t a c a a a c t	t g g a g t t c g c	t t g t a t t a c c	1500
g a g t a a t g g a	a t c c a t g c t t	a a a t c a g a a g	a a g a a c g a t t	a t c c a t t c a a	a a t t t t a g c a	1560
a a c t t c t g a a	t g a c a a c a t t	t t t c a t a t g t	c t t t a t t g g c	g t g c g c t c t t	g a g g t t g t a a	1620
t g g c c a c a t a	t a g c a g a a g t	a c a t c t c a g a	a t c t t g a t t c	t g g a a c a g a t	t t g t c t t t c c	1680
c a t g g a t t c t	g a a t g t g c t t	a a t t t a a a a g	c c t t t g a t t t	t t a c a a a g t g	a t c g a a a g t t	1740
t t a t c a a a g c	a g a a g g c a a c	t t g a c a a g a g	a a a t g a t a a a	a c a t t t a g a a	c g a t g t g a a c	1800
a t c g a a t c a t	g g a a t c c c t t	g c a t g g c t c t	c a g a t t c a c c	t t t a t t t g a t	c t t a t t a a a c	1860
a a t c a a a g g a	c c g a g a a g g a	c c a a c t g a t c	a c c t t g a a t c	t g c t t g t c c t	c t t a a t c t t c	1920
c t c t c c a g a a	t a a t c a c a c t	g c a g c a g a t a	t g t a t c t t t c	t c c t g t a a g a	t c t c c a a a g a	1980
a a a a a g g t t c	a a c t a c g c g t	g t a a a t t c t a	c t g c a a a t g c	a g a g a c a c a a	g c a a c c t c a g	2040
c c t t c c a g a c	c c a g a a g c c a	t t g a a a t c t a	c c t c t c t t t c	a c t g t t t t a t	a a a a a a g t g t	2100
a t c g g c t a g c	c t a t c t c c g g	c t a a a t a c a c	t t t g t g a a c g	c c t t c t g t c t	g a g c a c c c a g	2160
a a t t a g a a c a	t a t c a t c t g g	a c c c t t t t c c	a g c a c a c c c t	g c a g a a t g a g	t a t g a a c t c a	2220
t g a g a g a c a g	g c a t t t g g a c	c a a a t t a t g a	t g t g t t c c a t	g t a t g g c a t a	t g c a a a g t g a	2280
a g a a t a t a g a	c c t t a a a t t c	a a a a t c a t t g	t a a c a g c a t a	c a a g g a t c t t	c c t c a t g c t g	2340
t t c a g g a g a c	a t t c a a a c g t	g t t t t g a t c a	a a g a a g a g g a	g t a t g a t t c t	a t t a t a g t a t	2400
t c t a t a a c t c	g g t c t t c a t g	c a g a g a c t g a	a a a c a a a t a t	t t t g c a g t a t	g c t t c c a c c a	2460
g g c c c c c t a c	c t t g t c a c c a	a t a c c t c a c a	t t c c t c g a a g	c c c t t a c a a g	t t t c c t a g t t	2520
c a c c c t t a c g	g a t t c c t g g a	g g g a a c a t c t	a t a t t t c a c c	c c t g a a g a g t	c c a t a t a a a a	2580
t t t c a g a a g g	t c t g c c a a c a	c c a a c a a a a a	t g a c t c c a a g	a t c a a g a a t c	t t a g t a t c a a	2640
t t g g t g a a t c	a t t c g g g a c t	t c t g a g a a g t	t c c a g a a a a t	a a a t c a g a t g	g t a t g t a a c a	2700
g c g a c c g t g t	g c t c a a a a g a	a g t g c t g a a g	g a a g c a a c c c	t c c t a a a c c a	c t g a a a a a a c	2760
t a c g c t t t g a	t a t t g a a g g a	t c a g a t g a a g	c a g a t g g a a g	t a a a c a t c t c	c c a g g a g a g t	2820
c c a a a t t t c a	g c a g a a a c t g	g c a g a a a t g a	c t t c t a c t c g	a a c a c g a a t g	c a a a a g c a g a	2880
a a a t g a a t g a	t a g c a t g g a t	a c c t c a a a c a	a g g a a g a g a a	a t g a g g a t c t	c a g g a c c t t g	2940
g t g g a c a c t g	t g t a c a c c t c	t g g a t t c a t t	g t c t c t c a c a	g a t g t g a c t g	t a t a a c t t t c	3000
c c a g g t t c t g	t t t a t g g c c a	c a t t t a a t a t	c t t c a g c t c t	t t t t g t g g a t	a t a a a a t g t g	3060
c a g a t g c a a t	t g t t t g g g t g	a t t c c t a a g c	c a c t t g a a a t	g t t a g t c a t t	g t t a t t t a t a	3120
c a a g a t t g a a	a a t c t t g t g t	a a a t c c t g c c	a t t t a a a a a g	t t g t a g c a g a	t t g t t t c c t c	3180
t t c c a a a g t a	a a a t t g c t g t	g c t t t a t g g a	t a g t a a g a a t	g g c c c t a g a g	t g g g a g t c c t	3240
g a t a a c c c a g	g c c t g t c t g a	c t a c t t t g c c	t t c t t t t g t a	g c a t a t a g g t	g a t g t t t g c t	3300
c t t g t t t t t a	t t a a t t t t a t a	t g t a t a t t t t	t t t a a t t t a a	c a t g a a c a c c	c t t a g a a a a t	3360
g t g t c c t a t c	t a t c t t c c a a	a t g c a a t t t g	a t t g a c t g c c	c a t t c a c c a a	a a t t a t c c t g	3420
a a c t c t t c t g	c a a a a a t g g a	t a t t a t t a g a	a a t t a g a a a a	a a a t t a c t a a	t t t t a c a c a t	3480
t a g a t t t t a t	t t t a c t a t t g	g a a t c t g a t a	t a c t g t g t g c	t t g t t t t a t a	a a a t t t t g c t	3540
t t t a a t t a a a	t a a a a g c t g g	a a g c a a a g t a	t a a c c a t a t g	a t a c t a t c a t	a c t a c t g a a a	3600
c a g a t t t c a t	a c c t c a g a a t	g t a a a a g a a c	t t a c t g a t t a	t t t t c t t c a t	c c a a c t t a t g	3660
t t t t t a a a t g	a g g a t t a t t g	a t a g t a c t c t	t g g t t t t t a t	a c c a t t c a g a	t c a c t g a a t t	3720
t a t a a a g t a c	c c a t c t a g t a	c t t g a a a a a g	t a a a g t g t t c	t g c c a g a t c t	t a g g t a t a g a	3780
g g a c c c t a a c	a c a g t a t a t c	c c a a g t g c a c	t t t c t a a t g t	t t c t g g g t c c	t g a a g a a t t a	3840
a g a t a c a a a t	t a a t t t t a c t	c c a t a a a c a g	a c t g t t a a t t	a t a g g a g c c t	t a a t t t t t t t	3900
t t c a t a g a g a	t t t g t c t a a t	t g c a t c t c a a	a a t t a t t c t g	c c c t c c t t a a	t t t g g g a a g g	3960
t t t g t g t t t t	c t c t g g a a t g	g t a c a t g t c t	t c c a t g t a t c	t t t t g a a c t g	g c a a t t g t c t	4020
a t t t a t c t t t	t a t t t t t t t a	a g t c a g t a t g	g t c t a a c a c t	g g c a t g t t c a	a a g c c a c a t t	4080
a t t t c t a g t c	c a a a a t t a c a	a g t a a t c a a g	g g t c a t t a t g	g g t t a g g c a t	t a a t g t t t c t	4140
a t c t g a t t t t	g t g c a a a a g c	t t c a a a t t a a	a a c a g c t g c a	t t a g a a a a a g	a g g c g c t t c t	4200
c c c c t c c c c t	a c a c c t a a a g	g t g t a t t t a a	a c t a t c t t g t	g t g a t t a a c t	t a t t t a g a g a	4260
t g c t g t a a c t	t a a a a t a g g g	g a t a t t t a a g	g t a g c t t c a g	c t a g c t t t t a	g g a a a a t c a c	4320
t t t g t c t a a c	t c a g a a t t a t	t t t t a a a a a g	a a a t c t g g t c	t t g t t a g a a a	a c a a a a t t t t	4380
a t t t t g t g c t	c a t t t a a g t t	t c a a a c t t a c	t a t t t t g a c a	g t t a t t t t g a	t a a c a a t g a c	4440

actagaaaac	ttgactccat	ttcatcattg	tttctgcatg	aatatcatac	aaatcagtta	4500
gttttttaggt	caagggctta	ctatttctgg	gtcttttgct	actaagttca	cattagaatt	4560
agtgccagaa	tttttaggaac	ttcagagatc	gtgtattgag	atttcttaaa	taatgcttca	4620
gatattattg	ctttattgct	tttttgattt	ggttaaaact	gtacatttaa	aattgctatg	4680
ttactatttt	ctacaattaa	tagtttgtct	attttaaaat	aaattagttg	ttaagagtct	4740
taatgggtctg	atggtgtgtt	ctttgtatta	agtacactaa	tgttctcttt	tctgtctagg	4800
agaagataga	tagaagataa	ctctcctagt	atctcatcc			4839

<210> 9
 <211> 987
 <212> DNA
 <213> Homo sapiens

<400> 9		
cggagagggg	gagaacagac aacgggcggc ggggagcagc atggagccgg cggcggggag 60	
cagcatggag	ccttcggctg actggctggc cacggccggc gcccggggtc gggtagagga 120	
ggtgcgggcg	ctgctggagg cgggggcgct gcccaacgca ccgaatagtt acggtcggag 180	
gccgatccag	gtcatgatga tgggcagcgc ccgagtggcg gagctgctgc tgctccacgg 240	
cgcggagccc	aactgcgccg accccgccac tctcaccgca cccgtgcacg acgctgcccg 300	
ggagggttct	ctggacacgc tgggtggtgct gcaccggggc ggggcgcggc tggacgtgcg 360	
cgatgcctgg	ggccgtctgc ccgtggacct ggctgaggag ctggggccatc gcgatgtgcg 420	
acggtacctg	cgcgcggctg cggggggcac cagaggcagt aaccatgccc gcatagatgc 480	
cgcggaaggt	ccctcagaca tccccgattg aaagaaccag agaggctctg agaaacctcg 540	
ggaaacttag	atcatcagtc accgaaggct ctacagggcc acaactgccc ccgccacaac 600	
ccaccccgtc	ttcgtagttt tcatttagaa aatagagctt ttaaaaatgt cctgcctttt 660	
aacgtagata	taagccttcc cccactaccg taaatgtcca tttatatcat tttttatata 720	
ttcttataaa	aatgtaaaaa agaaaaacac cgcttctgcc ttttactgt gttggagttt 780	
tctggagtga	gcactcacgc cctaagcgca cattcatgtg ggcatctctt gcgagcctcg 840	
cagctccggg	aagctgtcga cttcatgaca agcattttgt gaactagggg agctcagggg 900	
ggttactggc	ttctcttgag tcacactgct agcaaatggc agaaccaaag ctcaataaaa 960	
aataaaaataa	ttttcattca ttcactc	987

<210> 10
 <211> 1181
 <212> DNA
 <213> Homo sapiens

<400> 10	
atggaggagc	cgcagtcaga tcctagcgtc gagccccctc tgagtcagga aacattttca 60
gacctatgga	aactacttcc tgaaaacaac gttctgtccc ccttgccgtc ccaagcaatg 120
gatgatttga	tgctgtcccc ggacgatatt gaacaatggt tcaactgaaga ccaggtcca 180
gatgaagctc	ccagaatgcc agaggctgct cccccgtgg cccctgcacc agcagctcct 240
acaccggcgg	cccctgcacc agccccctcc tggccccgtg catcttctgt cccttcccag 300
aaaacctacc	agggcagcta cggtttccgt ctgggcttct tgcattctgg gacagccaag 360
tctgtgactt	gcacgtactc ccctgccctc aacaagatgt tttgccaact ggccaagacc 420
tgccctgtgc	agctgtgggt tgattccaca cccccgcccg gcacccgcgt ccgcgccatg 480
gccatctaca	agcagtcaca gcacatgacg gaggttgtga ggcgtgccc ccacatgag 540
cgctgtcag	atagcgatgg tctggccccct cctcagcatc ttatccgagt ggaaggaaat 600
ttgcgtgtgg	agtatttgga tgacagaaac acttttgcac atagtgtggg ggtgccctat 660
gagccgcctg	aggttggctc tgactgtacc accatccact acaactacat gtgtaacagt 720
tcctgcatgg	gcggcatgaa ccggaggccc atcctcacca tcatcacact ggaagactcc 780
agtggtaatc	tactgggacg gaacagcttt gaggtgctgt tttgtgcctg tcctgggaga 840
gaccggcgca	cagaggaaga gaatctccgc aagaaagggg agcctcacca cgagctgccc 900
ccaggagca	ctaagcgagc actgcccac aacaccagct cctctcccca gccaaagaag 960
aaaccactgg	atggagaata tttcaccctt cagatccgtg ggcgtgagcg cttcgagatg 1020

```

ttccgagagc tgaatgaggc cttggaactc aaggatgccc aggctgggaa ggagccaggg 1080
gggagcaggg ctcactccag ccacctgaag tccaaaaagg gtcagtctac ctcccgccat 1140
aaaaaactca tgttcaagac agaaggccct gactcagact g 1181

```

```

<210> 11
<211> 540
<212> DNA
<213> Homo sapiens

```

```

<400> 11
cgcgccctgcg gggcgggagat gggcaggggg cggtgcgtgg gtcccagtct gcagttaagg 60
gggcaggagt ggcgctgctc acctctggtg ccaaagggcg gcgcagcggc tgccgagctc 120
ggccctggag gcgcgagaaa catggtgcgc aggttcttgg tgaccctccg gattcggcgc 180
gcgtgcggcc cgccgcgagt gagggttttc gtggttcaca tccgcgggt cacgggggag 240
tgggcagcgc caggggccc cgccgctgtg gccctcgtgc tgatgctact gaggagccag 300
cgtctagggc agcagccgct tcctagaaga ccaggtcatg atgatgggca gcgcccagat 360
ggcggagctg ctgctgctcc acggcgcgga gcccactgc gccgaccccg ccactctcac 420
ccgacccgtg cacgacgctg cccgggaggg cttcctggac acgctggtgg tgctgcaccg 480
ggccggggcg cggtggacg tgcgcgatgc ctggggccgt ctgcccgtgg acctggctga 540

```

```

<210> 12
<211> 26
<212> PRT
<213> Unknown Organism

```

```

<220>
<223> Description of Unknown Organism: internalizing
      peptide

```

```

<400> 12
Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly Arg Lys Lys Arg
  1             5             10             15
Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser
      20             25

```

```

<210> 13
<211> 12
<212> PRT
<213> Unknown Organism

```

```

<220>
<223> Description of Unknown Organism: EGF derived
      peptide

```

```

<400> 13
Cys Met His Ile Glu Ser Leu Asp Ser Tyr Thr Cys
  1             5             10

```

```

<210> 14
<211> 12
<212> PRT
<213> Unknown Organism

```

<220>
<223> Description of Unknown Organism: EGF derived
peptide

<400> 14
Cys Met Tyr Ile Glu Ala Leu Asp Lys Tyr Ala Cys
1 5 10

<210> 15
<211> 32
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: internalizing
peptide

<220>
<221> MOD_RES
<222> (1)
<223> unique residue, such as cysteine or lysine, that
facilitates chemical conjugation of the
internalizing peptide to a target protein
conjugate.

<220>
<221> MOD_RES
<222> (2)..(3)
<223> amino acid residue selected to modulate the
affinity of the internalizing peptide for
different membranes.

<400> 15
Xaa Xaa Xaa Glu Ala Ala Leu Ala Glu Ala Leu Ala Glu Ala Leu Ala
1 5 10 15

Glu Ala Leu Ala Glu Ala Leu Ala Glu Ala Leu Glu Ala Leu Ala Ala
20 25 30

<210> 16
<211> 8
<212> PRT
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: peptide substrate

<400> 16
Gly Asn Ala Ala Ala Arg Arg
1 5

<210> 17
<211> 10
<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: laminin derived peptide

<400> 17

Cys Asp Pro Gly Tyr Ile Gly Ser Arg Cys
1 5 10

<210> 18

<211> 75

<212> DNA

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: Nde1-EcoR1 fragment

<400> 18

catatgggtg gctgccgtgg cgatatgttc ggttgcggtg ctcctccaaa aaagaagaga 60
aaggtagctg gattc 75

<210> 19

<211> 24

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: RGD/SV40 peptide

<400> 19

Met Gly Gly Cys Arg Gly Asp Met Phe Gly Cys Gly Ala Pro Pro Lys
1 5 10 15

Lys Lys Arg Lys Val Ala Gly Phe
20

<210> 20

<211> 225

<212> DNA

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: Nde1-EcoR1 fragment

<400> 20

catatgggagc cagtagatcc tagactagag ccctggaagc atccaggaag tcagcctaaa 60
actgcttgta ccaattgcta ttgtaaaaag tggtgctttc attgccaaagt ttgtttcata 120
acaaaagccc ttggcatctc ctatggcagg aagaagcgga gacagcgacg aagacctcct 180
caaggcagtc agactcatca agtttctcta agtaagcaag gattc 225

<210> 21
<211> 72
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: engineered HIV-1 tat

<400> 21
Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser
1 5 10 15
Gln Pro Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe
20 25 30
His Cys Gln Val Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly
35 40 45
Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Gly Ser Gln Thr
50 55 60
His Gln Val Ser Leu Ser Lys Gln
65 70

<210> 22
<211> 912
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: NdeI-EcoRI
fragment

<400> 22
catatgacct ctccgctc cgtgaagtcg ggtccgcggg aggttccgcg cgatgagtac 60
gaggatctgt actacacccc gtcttcaggt atggcgagtc ccgatagtc gcctgacacc 120
tcccgcctg gcgcctaca gacacgctcg cgccagaggg gcgaggtccg tttcgtccag 180
tacgacgagt cggattatgc cctctacggg ggctcgtcat ccgaagacga cgaacacccg 240
gaggtccccc ggacgcggcg tcccgtttcc ggggcggttt tgtccggccc ggggcctgcg 300
cgggcgctc cgccaccgc tgggtccgga ggggcggac gcacaccac caccgcccc 360
cgggccccc gaaccagcg ggtggcgact aaggccccc cgccccgcg gcgagagacc 420
acccgcggca ggaaatcggc ccagccagaa tccgcgcac tcccagacgc cccgcgctcg 480
acggcgccaa cccgatccaa gacaccgcg caggggctgg ccagaaagct gcacttttagc 540
accgcccc caaacccga cgcgccatgg acccccggg tggccggctt taacaagcgc 600
gtcttctgcg ccgcggtcgg gcgectggcg gccatgcatg ccggtatggc ggcggtccag 660
ctctgggaca tgtcgcgtcc gcgcacagac gaagacctca acgaactcct tggcatcacc 720
accatccgcg tgacggtctg cgagggcaaa aacctgcttc agcgcgccaa cgagttggtg 780
aatccagacg tgggtcgagga cgtcgacgcg gccacggcga ctcgagggcg ttctgcggcg 840
tcgcgcccc cagagcgacc tcgagcccc gcccgctccg cttctcgccc cagacggccc 900
gtcgaggaat tc 912

<210> 23
<211> 301
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: engineered HSV-1 VP22

<400> 23

Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15

Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
35 40 45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
50 55 60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
65 70 75 80

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
85 90 95

Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
100 105 110

Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gly Arg Val Ala
115 120 125

Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
130 135 140

Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145 150 155 160

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
165 170 175

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
180 185 190

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
195 200 205

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
210 215 220

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225 230 235 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
245 250 255

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
260 265 270

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala

275

280

285

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu
290 295 300

<210> 24
<211> 120
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Nde1-EcoR1
fragment

<400> 24
catatggacg tgcacgcggc cacggcgact cgagggcggt ctgcggcgct gcgccccacc 60
gagcgacctc gagccccagc ccgctccgct tctcgcccca gacggcccgt cgaggaattc 120

<210> 25
<211> 37
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: VP22 (C-terminal domain)

<400> 25
Met Asp Val Asp Ala Ala Thr Ala Thr Arg Gly Arg Ser Ala Ala Ser
1 5 10 15
Arg Pro Thr Glu Arg Pro Arg Ala Pro Ala Arg Ser Ala Ser Arg Pro
20 25 30
Arg Arg Pro Val Glu
35

<210> 26
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: illustrative.
peptide

<400> 26
Cys Ala Ala Met
1

<210> 27
<211> 4
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: illustrative peptide

<400> 27

Cys Ala Ala Ser

1

<210> 28

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 28

aatccgtcga gcagagtt

18

<210> 29

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 29

gcgcggctaa ccctaaccct aacc

24

<210> 30

<211> 1031

<212> PRT

<213> Euplotes sp.

<400> 30

Met Glu Val Asp Val Asp Asn Gln Ala Asp Asn His Gly Ile His Ser
1 5 10 15

Ala Leu Lys Thr Cys Glu Glu Ile Lys Glu Ala Lys Thr Leu Tyr Ser
20 25 30

Asn Ile Gln Lys Val Ile Arg Cys Arg Asn Gln Ser Gln Ser His Tyr
35 40 45

Lys Asp Leu Glu Asp Ile Lys Ile Phe Ala Gln Thr Asn Ile Val Ala
50 55 60

Thr Pro Arg Asp Tyr Asn Glu Glu Asp Phe Lys Val Ile Ala Arg Lys
65 70 75 80

Glu Val Phe Ser Thr Gly Leu Met Ile Glu Leu Ile Asp Lys Cys Leu
85 90 95

Val	Glu	Leu	Leu	Ser	Ser	Ser	Asp	Val	Ser	Asp	Arg	Gln	Lys	Leu	Gln	100	105	110
Cys	Phe	Gly	Phe	Gln	Leu	Lys	Gly	Asn	Gln	Leu	Ala	Lys	Thr	His	Leu	115	120	125
Leu	Thr	Ala	Leu	Ser	Thr	Gln	Lys	Gln	Tyr	Phe	Phe	Gln	Asp	Glu	Trp	130	135	140
Asn	Gln	Val	Arg	Ala	Met	Ile	Gly	Asn	Glu	Leu	Phe	Arg	His	Leu	Tyr	145	150	155
Thr	Lys	Tyr	Leu	Ile	Phe	Gln	Arg	Thr	Ser	Glu	Gly	Thr	Leu	Val	Gln	165	170	175
Phe	Cys	Gly	Asn	Asn	Val	Phe	Asp	His	Leu	Lys	Val	Asn	Asp	Lys	Phe	180	185	190
Asp	Lys	Lys	Gln	Lys	Gly	Gly	Ala	Ala	Asp	Met	Asn	Glu	Pro	Arg	Cys	195	200	205
Cys	Ser	Thr	Cys	Lys	Tyr	Asn	Val	Lys	Asn	Glu	Lys	Asp	His	Phe	Leu	210	215	220
Asn	Asn	Ile	Asn	Val	Pro	Asn	Asn	Asn	Asn	Met	Lys	Ser	Arg	Thr	Arg	225	230	235
Ile	Phe	Tyr	Cys	Thr	His	Phe	Asn	Arg	Asn	Asn	Gln	Phe	Phe	Lys	Lys	245	250	255
His	Glu	Phe	Val	Ser	Asn	Lys	Asn	Asn	Ile	Ser	Ala	Met	Asp	Arg	Ala	260	265	270
Gln	Thr	Ile	Phe	Thr	Asn	Ile	Phe	Arg	Phe	Asn	Arg	Ile	Arg	Lys	Lys	275	280	285
Leu	Lys	Asp	Lys	Val	Ile	Glu	Lys	Ile	Ala	Tyr	Met	Leu	Glu	Lys	Val	290	295	300
Lys	Asp	Phe	Asn	Phe	Asn	Tyr	Tyr	Leu	Thr	Lys	Ser	Cys	Pro	Leu	Pro	305	310	315
Glu	Asn	Asn	Arg	Glu	Arg	Lys	Gln	Lys	Ile	Glu	Asn	Leu	Ile	Asn	Lys	325	330	335
Thr	Arg	Glu	Glu	Lys	Ser	Lys	Tyr	Tyr	Glu	Glu	Leu	Phe	Ser	Tyr	Thr	340	345	350
Thr	Asp	Asn	Lys	Cys	Val	Thr	Gln	Phe	Ile	Asn	Glu	Phe	Phe	Tyr	Asn	355	360	365
Ile	Leu	Pro	Lys	Asp	Phe	Leu	Thr	Gly	Arg	Asn	Arg	Lys	Asn	Phe	Gln	370	375	380
Lys	Lys	Val	Lys	Lys	Tyr	Val	Glu	Leu	Asn	Lys	His	Glu	Leu	Ile	His	385	390	395
																		400

Lys Asn Leu Leu Leu Glu Lys Ile Asn Thr Arg Glu Ile Ser Trp Met
 405 410 415
 Gln Val Glu Thr Ser Ala Lys His Phe Tyr Tyr Phe Asp His Glu Asn
 420 425 430
 Ile Tyr Val Leu Trp Lys Leu Leu Arg Trp Ile Phe Glu Asp Leu Val
 435 440 445
 Val Ser Leu Ile Arg Cys Phe Phe Tyr Val Thr Glu Gln Gln Lys Ser
 450 455 460
 Tyr Ser Lys Thr Tyr Tyr Tyr Arg Lys Asn Ile Trp Asp Val Ile Met
 465 470 475 480
 Lys Met Ser Ile Ala Asp Leu Lys Lys Glu Thr Leu Ala Glu Val Gln
 485 490 495
 Glu Lys Glu Val Glu Glu Trp Lys Lys Ser Leu Gly Phe Ala Pro Gly
 500 505 510
 Lys Leu Arg Leu Ile Pro Lys Lys Thr Thr Phe Arg Pro Ile Met Thr
 515 520 525
 Phe Asn Lys Lys Ile Val Asn Ser Asp Arg Lys Thr Thr Lys Leu Thr
 530 535 540
 Thr Asn Thr Lys Leu Leu Asn Ser His Leu Met Leu Lys Thr Leu Lys
 545 550 555 560
 Asn Arg Met Phe Lys Asp Pro Phe Gly Phe Ala Val Phe Asn Tyr Asp
 565 570 575
 Asp Val Met Lys Lys Tyr Glu Glu Phe Val Cys Lys Trp Lys Gln Val
 580 585 590
 Gly Gln Pro Lys Leu Phe Phe Ala Thr Met Asp Ile Glu Lys Cys Tyr
 595 600 605
 Asp Ser Val Asn Arg Glu Lys Leu Ser Thr Phe Leu Lys Thr Thr Lys
 610 615 620
 Leu Leu Ser Ser Asp Phe Trp Ile Met Thr Ala Gln Ile Leu Lys Arg
 625 630 635 640
 Lys Asn Asn Ile Val Ile Asp Ser Lys Asn Phe Arg Lys Lys Glu Met
 645 650 655
 Lys Asp Tyr Phe Arg Gln Lys Phe Gln Lys Ile Ala Leu Glu Gly Gly
 660 665 670
 Gln Tyr Pro Thr Leu Phe Ser Val Leu Glu Asn Glu Gln Asn Asp Leu
 675 680 685
 Asn Ala Lys Lys Thr Leu Ile Val Glu Ala Lys Gln Arg Asn Tyr Phe
 690 695 700

Lys Lys Asp Asn Leu Leu Gln Pro Val Ile Asn Ile Cys Gln Tyr Asn
 705 710 715 720
 Tyr Ile Asn Phe Asn Gly Lys Phe Tyr Lys Gln Thr Lys Gly Ile Pro
 725 730 735
 Gln Gly Leu Cys Val Ser Ser Ile Leu Ser Ser Phe Tyr Tyr Ala Thr
 740 745 750
 Leu Glu Glu Ser Ser Leu Gly Phe Leu Arg Asp Glu Ser Met Asn Pro
 755 760 765
 Glu Asn Pro Asn Val Asn Leu Leu Met Arg Leu Thr Asp Asp Tyr Leu
 770 775 780
 Leu Ile Thr Thr Gln Glu Asn Asn Ala Val Leu Phe Ile Glu Lys Leu
 785 790 795 800
 Ile Asn Val Ser Arg Glu Asn Gly Phe Lys Phe Asn Met Lys Lys Leu
 805 810 815
 Gln Thr Ser Phe Pro Leu Ser Pro Ser Lys Phe Ala Lys Tyr Gly Met
 820 825 830
 Asp Ser Val Glu Glu Gln Asn Ile Val Gln Asp Tyr Cys Asp Trp Ile
 835 840 845
 Gly Ile Ser Ile Asp Met Lys Thr Leu Ala Leu Met Pro Asn Ile Asn
 850 855 860
 Leu Arg Ile Glu Gly Ile Leu Cys Thr Leu Asn Leu Asn Met Gln Thr
 865 870 875 880
 Lys Lys Ala Ser Met Trp Leu Lys Lys Lys Leu Lys Ser Phe Leu Met
 885 890 895
 Asn Asn Ile Thr His Tyr Phe Arg Lys Thr Ile Thr Thr Glu Asp Phe
 900 905 910
 Ala Asn Lys Thr Leu Asn Lys Leu Phe Ile Ser Gly Gly Tyr Lys Tyr
 915 920 925
 Met Gln Cys Ala Lys Glu Tyr Lys Asp His Phe Lys Lys Asn Leu Ala
 930 935 940
 Met Ser Ser Met Ile Asp Leu Glu Val Ser Lys Ile Ile Tyr Ser Val
 945 950 955 960
 Thr Arg Ala Phe Phe Lys Tyr Leu Val Cys Asn Ile Lys Asp Thr Ile
 965 970 975
 Phe Gly Glu Glu His Tyr Pro Asp Phe Phe Leu Ser Thr Leu Lys His
 980 985 990
 Phe Ile Glu Ile Phe Ser Thr Lys Lys Tyr Ile Phe Asn Arg Val Cys
 995 1000 1005

Met Ile Leu Lys Ala Lys Glu Ala Lys Leu Lys Ser Asp Gln Cys Gln
1010 1015 1020

Ser Leu Ile Gln Tyr Asp Ala
1025 1030

<210> 31

<211> 884

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 31

Met Lys Ile Leu Phe Glu Phe Ile Gln Asp Lys Leu Asp Ile Asp Leu
1 5 10 15

Gln Thr Asn Ser Thr Tyr Lys Glu Asn Leu Lys Cys Gly His Phe Asn
20 25 30

Gly Leu Asp Glu Ile Leu Thr Thr Cys Phe Ala Leu Pro Asn Ser Arg
35 40 45

Lys Ile Ala Leu Pro Cys Leu Pro Gly Asp Leu Ser His Lys Ala Val
50 55 60

Ile Asp His Cys Ile Ile Tyr Leu Leu Thr Gly Glu Leu Tyr Asn Asn
65 70 75 80

Val Leu Thr Phe Gly Tyr Lys Ile Ala Arg Asn Glu Asp Val Asn Asn
85 90 95

Ser Leu Phe Cys His Ser Ala Asn Val Asn Val Thr Leu Leu Lys Gly
100 105 110

Ala Ala Trp Lys Met Phe His Ser Leu Val Gly Thr Tyr Ala Phe Val
115 120 125

Asp Leu Leu Ile Asn Tyr Thr Val Ile Gln Phe Asn Gly Gln Phe Phe
130 135 140

Thr Gln Ile Val Gly Asn Arg Cys Asn Glu Pro His Leu Pro Pro Lys
145 150 155 160

Trp Val Gln Arg Ser Ser Ser Ser Ser Ala Thr Ala Ala Gln Ile Lys
165 170 175

Trp Leu Thr Glu Pro Val Thr Asn Lys Gln Phe Leu His Lys Leu Asn
180 185 190

Ile Asn Ser Ser Ser Phe Phe Pro Tyr Ser Lys Ile Leu Pro Ser Ser
195 200 205

Ser Ser Ile Lys Lys Leu Thr Asp Leu Arg Glu Ala Ile Phe Pro Thr
210 215 220

Asn Leu Val Lys Ile Pro Gln Arg Leu Lys Val Arg Ile Asn Leu Thr

225	230										235					240				
Leu	Gln	Lys	Leu	Leu	Lys	Arg	His	Lys	Arg	Leu	Asn	Tyr	Val	Ser	Ile					
				245					250					255						
Leu	Asn	Ser	Ile	Cys	Pro	Pro	Leu	Glu	Gly	Thr	Val	Leu	Asp	Leu	Ser					
			260					265					270							
His	Leu	Ser	Arg	Gln	Ser	Pro	Lys	Glu	Arg	Val	Leu	Lys	Phe	Ile	Ile					
		275					280					285								
Val	Ile	Leu	Gln	Lys	Leu	Leu	Pro	Gln	Glu	Met	Phe	Gly	Ser	Lys	Lys					
	290					295					300									
Asn	Lys	Gly	Lys	Ile	Ile	Lys	Asn	Leu	Asn	Leu	Leu	Leu	Ser	Leu	Pro					
305					310					315					320					
Leu	Asn	Gly	Tyr	Leu	Pro	Phe	Asp	Ser	Leu	Leu	Lys	Lys	Leu	Arg	Leu					
				325					330					335						
Lys	Asp	Phe	Arg	Trp	Leu	Phe	Ile	Ser	Asp	Ile	Trp	Phe	Thr	Lys	His					
			340					345					350							
Asn	Phe	Glu	Asn	Leu	Asn	Gln	Leu	Ala	Ile	Cys	Phe	Ile	Ser	Trp	Leu					
		355					360					365								
Phe	Arg	Gln	Leu	Ile	Pro	Lys	Ile	Ile	Gln	Thr	Phe	Phe	Tyr	Cys	Thr					
	370					375					380									
Glu	Ile	Ser	Ser	Thr	Val	Thr	Ile	Val	Tyr	Phe	Arg	His	Asp	Thr	Trp					
385					390					395					400					
Asn	Lys	Leu	Ile	Thr	Pro	Phe	Ile	Val	Glu	Tyr	Phe	Lys	Thr	Tyr	Leu					
				405					410					415						
Val	Glu	Asn	Asn	Val	Cys	Arg	Asn	His	Asn	Ser	Tyr	Thr	Leu	Ser	Asn					
			420					425					430							
Phe	Asn	His	Ser	Lys	Met	Arg	Ile	Ile	Pro	Lys	Lys	Ser	Asn	Asn	Glu					
		435					440					445								
Phe	Arg	Ile	Ile	Ala	Ile	Pro	Cys	Arg	Gly	Ala	Asp	Glu	Glu	Glu	Phe					
	450					455					460									
Thr	Ile	Tyr	Lys	Glu	Asn	His	Lys	Asn	Ala	Ile	Gln	Pro	Thr	Gln	Lys					
465					470					475					480					
Ile	Leu	Glu	Tyr	Leu	Arg	Asn	Lys	Arg	Pro	Thr	Ser	Phe	Thr	Lys	Ile					
				485					490					495						
Tyr	Ser	Pro	Thr	Gln	Ile	Ala	Asp	Arg	Ile	Lys	Glu	Phe	Lys	Gln	Arg					
			500					505					510							
Leu	Leu	Lys	Lys	Phe	Asn	Asn	Val	Leu	Pro	Glu	Leu	Tyr	Phe	Met	Lys					
		515					520					525								
Phe	Asp	Val	Lys	Ser	Cys	Tyr	Asp	Ser	Ile	Pro	Arg	Met	Glu	Cys	Met					

530	535	540
Arg Ile Leu Lys Asp 545	Ala Leu Lys Asn Glu 550	Asn Gly Phe Phe Val Arg 555 560
Ser Gln Tyr Phe 565	Phe Asn Thr Asn Thr Gly Val 570	Leu Lys Leu Phe Asn 575
Val Val Asn Ala Ser Arg Val Pro Lys Pro Tyr Glu Leu Tyr Ile Asp 580 585 590		
Asn Val Arg Thr Val His Leu Ser Asn Gln Asp Val Ile Asn Val Val 595 600 605		
Glu Met Glu Ile Phe Lys Thr Ala Leu Trp Val Glu Asp Lys Cys Tyr 610 615 620		
Ile Arg Glu Asp Gly Leu Phe Gln Gly Ser Ser Leu Ser Ala Pro Ile 625 630 635 640		
Val Asp Leu Val Tyr Asp Asp Leu Leu Glu Phe Tyr Ser Glu Phe Lys 645 650 655		
Ala Ser Pro Ser Gln Asp Thr Leu Ile Leu Lys Leu Ala Asp Asp Phe 660 665 670		
Leu Ile Ile Ser Thr Asp Gln Gln Gln Val Ile Asn Ile Lys Lys Leu 675 680 685		
Ala Met Gly Gly Phe Gln Lys Tyr Asn Ala Lys Ala Asn Arg Asp Lys 690 695 700		
Ile Leu Ala Val Ser Ser Gln Ser Asp Asp Asp Thr Val Ile Gln Phe 705 710 715 720		
Cys Ala Met His Ile Phe Val Lys Glu Leu Glu Val Trp Lys His Ser 725 730 735		
Ser Thr Met Asn Asn Phe His Ile Arg Ser Lys Ser Ser Lys Gly Ile 740 745 750		
Phe Arg Ser Leu Ile Ala Leu Phe Asn Thr Arg Ile Ser Tyr Lys Thr 755 760 765		
Ile Asp Thr Asn Leu Asn Ser Thr Asn Thr Val Leu Met Gln Ile Asp 770 775 780		
His Val Val Lys Asn Ile Ser Glu Cys Tyr Lys Ser Ala Phe Lys Asp 785 790 795 800		
Leu Ser Ile Asn Val Thr Gln Asn Met Gln Phe His Ser Phe Leu Gln 805 810 815		
Arg Ile Ile Glu Met Thr Val Ser Gly Cys Pro Ile Thr Lys Cys Asp 820 825 830		
Pro Leu Ile Glu Tyr Glu Val Arg Phe Thr Ile Leu Asn Gly Phe Leu		

835

840

845

Glu Ser Leu Ser Ser Asn Thr Ser Lys Phe Lys Asp Asn Ile Ile Leu
 850 855 860

Leu Arg Lys Glu Ile Gln His Leu Gln Ala Tyr Ile Tyr Ile Tyr Ile
 865 870 875 880

His Ile Val Asn

<210> 32

<211> 8

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 32

Val Arg Arg Val Glu Ile Pro Lys
 1 5

<210> 33

<211> 12

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 33

Phe Arg Pro Leu Ser Val Gly Asn Pro Arg Glu Lys
 1 5 10

<210> 34

<211> 17

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 34

Tyr Cys Asn Trp Phe Ile Lys Val Asp Leu Asn Lys Cys Phe Asp Thr
 1 5 10 15

Ile

<210> 35

<211> 24

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 35

Asn Tyr His Asn Thr Thr Leu Gly Ile Pro Gln Gly Ser Val Val Ser
 1 5 10 15

Pro Ile Leu Cys Asn Ile Phe Leu
 20

<210> 36
<211> 13
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 36
Tyr Phe Val Arg Tyr Ala Asp Asp Ile Ile Ile Gly Val
1 5 10

<210> 37
<211> 11
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 37
Gly Met Ser Ile Asn Ile Asp Lys Ser Val Ile
1 5 10

<210> 38
<211> 8
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 38
Met Arg Ile Val Asn Ile Pro Lys
1 5

<210> 39
<211> 12
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 39
Ile Arg Pro Leu Ser Val Gly Asn Pro Arg Asp Lys
1 5 10

<210> 40
<211> 17
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 40
Gly Ser Asn Trp Phe Arg Glu Val Asp Leu Lys Lys Cys Phe Asp Thr
1 5 10 15

Ile

<210> 41
<211> 25
<212> PRT

<213> Saccharomyces cerevisiae

<400> 41

Thr Tyr His Lys Pro Ile Leu Gly Leu Pro Gln Gly Ser Leu Ile Pro
1 5 10 15

Ser Pro Ile Leu Cys Asn Ile Val Ile
20 25

<210> 42

<211> 13

<212> PRT

<213> Saccharomyces cerevisiae

<400> 42

Lys Tyr Val Arg Tyr Ala Asp Asp Ile Leu Ile Gly Val
1 5 10

<210> 43

<211> 11

<212> PRT

<213> Saccharomyces cerevisiae

<400> 43

Gly Leu Thr Ile Asn Glu Glu Lys Thr Leu Ile
1 5 10